Adaptation, Learning, and Optimization over Networks

Ali H. Sayed
University of California at Los Angeles

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Machine Learning
Vol. 7, No. 4-5 (2014) 311–801
© 2014 A. H. Sayed
DOI: 10.1561/2200000051

now
the essence of knowledge

Adaptation, Learning, and Optimization over Networks

Ali H. Sayed
University of California at Los Angeles

Contents

1 Motivation and Notation **3**

1.1 Introduction 3
1.2 Biological Networks 4
1.3 Distributed Processing 4
1.4 Adaptive Networks 6
1.5 Organization 7
1.6 Notation and Symbols 8

2 Optimization by Single Agents **11**

2.1 Risk and Loss Functions 11
2.2 Conditions on Risk Function 15
2.3 Optimization via Gradient Descent 17
2.4 Decaying Step-Size Sequences 20
2.5 Optimization in the Complex Domain 25

3 Stochastic Optimization by Single Agents **31**

3.1 Adaptation and Learning 32
3.2 Gradient Noise Process 36
3.3 Stability of Second-Order Error Moment 39
3.4 Stability of Fourth-Order Error Moment 42
3.5 Decaying Step-Size Sequences 48

3.6 Optimization in the Complex Domain 52

4 Performance of Single Agents 61
4.1 Conditions on Risk Function and Noise 63
4.2 Stability of First-Order Error Moment 70
4.3 Long-Term Error Dynamics 72
4.4 Size of Approximation Error 76
4.5 Performance Metrics 79
4.6 Performance in the Complex Domain 93

5 Centralized Adaptation and Learning 101
5.1 Non-Cooperative Processing 101
5.2 Centralized Processing 105
5.3 Stochastic-Gradient Centralized Solution 107
5.4 Gradient Noise Model 109
5.5 Performance of Centralized Solution 113
5.6 Comparison with Single Agents 116
5.7 Decaying Step-Size Sequences 123

6 Multi-Agent Network Model 125
6.1 Connected Networks 125
6.2 Strongly-Connected Networks 129
6.3 Network Objective 134

7 Multi-Agent Distributed Strategies 143
7.1 Incremental Strategy 144
7.2 Consensus Strategy 147
7.3 Diffusion Strategy 151

8 Evolution of Multi-Agent Networks 165
8.1 State Recursion for Network Errors 165
8.2 Network Limit Point and Pareto Optimality 173
8.3 Gradient Noise Model 191
8.4 Extended Network Error Dynamics 193

9 Stability of Multi-Agent Networks 203
9.1 Stability of Second-Order Error Moment 204

9.2 Stability of Fourth-Order Error Moment 218
9.3 Stability of First-Order Error Moment 227

10 Long-Term Network Dynamics **249**
10.1 Long-Term Error Model 250
10.2 Size of Approximation Error 253
10.3 Stability of Second-Order Error Moment 257
10.4 Stability of Fourth-Order Error Moment 260
10.5 Stability of First-Order Error Moment 263
10.6 Comparing Consensus and Diffusion Strategies 265

11 Performance of Multi-Agent Networks **271**
11.1 Conditions on Costs and Noise 272
11.2 Performance Metrics 278
11.3 Mean-Square-Error Performance 280
11.4 Excess-Risk Performance 305
11.5 Comparing Consensus and Diffusion Strategies 312

12 Benefits of Cooperation **321**
12.1 Doubly-Stochastic Combination Policies 322
12.2 Left-Stochastic Combination Policies 325
12.3 Comparison with Centralized Solutions 332
12.4 Excess-Risk Performance 338

13 Role of Informed Agents **343**
13.1 Informed and Uninformed Agents 343
13.2 Conditions on Cost Functions 345
13.3 Mean-Square-Error Performance 347
13.4 Controlling Degradation in Performance 356
13.5 Excess-Risk Performance 357

14 Combination Policies **359**
14.1 Static Combination Policies 360
14.2 Need for Adaptive Policies 362
14.3 Hastings Policy 364
14.4 Relative-Variance Policy 365

14.5 Adaptive Combination Policy 368

15 Extensions and Conclusions **381**
15.1 Gossip and Asynchronous Strategies 381
15.2 Noisy Exchanges of Information 384
15.3 Exploiting Temporal Diversity 385
15.4 Incorporating Sparsity Constraints 388
15.5 Distributed Constrained Optimization 389
15.6 Distributed Recursive Least-Squares 394
15.7 Distributed State-Space Estimation 399

Acknowledgements **408**

Appendices **411**

A Complex Gradient Vectors **413**
A.1 Cauchy-Riemann Conditions 413
A.2 Scalar Arguments 415
A.3 Vector Arguments 417
A.4 Real Arguments 419

B Complex Hessian Matrices **421**
B.1 Hessian Matrices for Real Arguments 421
B.2 Hessian Matrices for Complex Arguments 423

C Convex Functions **431**
C.1 Convexity in the Real Domain 432
C.2 Convexity in the Complex Domain 440

D Mean-Value Theorems **445**
D.1 Increment Formulae for Real Arguments 445
D.2 Increment Formulae for Complex Arguments 447

E Lipschitz Conditions **451**
E.1 Perturbation Bounds in the Real Domain 451
E.2 Lipschitz Conditions in the Real Domain 455

E.3 Perturbation Bounds in the Complex Domain 457
E.4 Lipschitz Conditions in the Complex Domain 461

F Useful Matrix and Convergence Results **463**
F.1 Kronecker Products 463
F.2 Vector and Matrix Norms 466
F.3 Perturbation Bounds on Eigenvalues 472
F.4 Lyapunov Equations 474
F.5 Stochastic Matrices 476
F.6 Convergence of Inequality Recursions 477

G Logistic Regression **479**
G.1 Logistic Function 479
G.2 Odds Function 480
G.3 Kullback-Leibler Divergence 481

References **483**